PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 3 :
C07H 17/00, 19/06, 15/12
C12Q 1/68

A1

(11) International Publication Number: WO 84/ 03285

(43) International Publication Date: 30 August 1984 (30.08.84)

(21) International Application Number: PCT/US84/00279

(22) International Filing Date: 22 February 1984 (22.02.84)

(31) Priority Application Numbers: PCT/US83/00254
468,498

(32) Priority Dates: 22 February 1983 (22.02.83)
22 February 1983 (22.02.83)

(33) Priority Countries: AU
US

(60) Parent Application or Grant
(63) Related by Continuation
US 468,498 (CIP)
Filed on 22 February 1983 (22.02.83)

(71) Applicant *(for all designated States except US)*: MOLECULAR BIOSYSTEMS, INC. [US/US]; 11180 Roselle Street, Suite A, San Diego, CA 92121 (US).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)* : RUTH, Jerry, L. [US/US]; 2433 Levante Street, Carlsbad, CA 92008 (US).

(74) Agent: SCHWALBACH, Joseph, C.; 1010 Second Avenue, Suite 1607, San Diego, CA 92101 (US).

(81) Designated States: AU, CH (European patent), DE (European patent), DK, FR (European patent), GB (European patent), JP, NL (European patent), NO, SE (European patent), US.

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: DEFINED SEQUENCE SINGLE STRAND OLIGONUCLEOTIDES INCORPORATING REPORTER GROUPS, PROCESS FOR THE CHEMICAL SYNTHESIS THEREOF, AND NUCLEOSIDES USEFUL IN SUCH SYNTHESIS

(57) Abstract

Defined sequence single strand oligonucleotides which have a length of fewer than 200 units, incorporate modified nucleotide units which are pyrimidine or purine-based, and which may have readily detectable reporter groups attached to substituents on the modified nucleotide units thereof. Each reporter group is attached to the base of the respective nucleotide unit at a sterically tolerant site thereon, exemplified by the C-5 position of pyrimidine-based nucleotides and the C-8 position of purine-based nucleotides. Such oligonucleotides are useful in identification, localization, isolation and/or detection of complementary nucleic acid sequences of interest in cellular or cell-free systems. Also disclosed is a process for the chemical synthesis of single strand oligonucleotides having a predetermined number of modified pyrimidine and/or purine-based nucleotide units in predetermined sequence, to the bases of which readily detectable reporter groups are attached at sterically tolerant sites through substituents thereon, either prior to or after incorporation of such units into the oligonucleotide chain. Novel nucleosides useful in the chemical synthesis of labeled, defined sequence single strand oligonucleotides are also disclosed.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**GA**	Gabon	**MR**	Mauritania
AU	Australia	**GB**	United Kingdom	**MW**	Malawi
BB	Barbados	**HU**	Hungary	**NL**	Netherlands
BE	Belgium	**IT**	Italy	**NO**	Norway
BG	Bulgaria	**JP**	Japan	**RO**	Romania
BR	Brazil	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CF	Central African Republic			**SE**	Sweden
CG	Congo	**KR**	Republic of Korea	**SN**	Senegal
CH	Switzerland	**LI**	Liechtenstein	**SU**	Soviet Union
CM	Cameroon	**LK**	Sri Lanka	**TD**	Chad
DE	Germany, Federal Republic of	**LU**	Luxembourg	**TG**	Togo
DK	Denmark	**MC**	Monaco	**US**	United States of America
FI	Finland	**MG**	Madagascar		
FR	France	**ML**	Mali		

Description

Defined Sequence Single Strand Oligonucleotides Incorporating Reporter Groups, Process for the Chemical Synthesis Thereof, and Nucleosides Useful in Such Synthesis

Technical Field

The present invention relates to defined sequence single strand oligonucleotides having a length of fewer than 200 base units and which contain one or more nucleotide units to the base of which is attached a substituent group capable of functioning as one or more detectable reporter groups, or which has bound thereto one or more detectable reporter groups; such oligonucleotides being useful in the identification, localization and detection of complementary nucleic acid sequences of interest in cellular or cell-free systems.

Background Art

The enzymatic production of labelled, double stranded deoxypolynucleotides has been accomplished with prior art techniques involving the incorporation of radioisotopes into double-strand DNA by the nick translation protocol of P. Rigby et al, J. Mol. Biol. 113: 237-251 (1977), or the gap-filling reaction described by G. Bourguignon et al, J. Virol. 20: 290-306 (1976). Specifically, a nick is introduced via DNase, and then translated along the DNA strand using a DNA polymerase. During the nick translation procedure, DNA polymerase from E. coli (Pol I) will, in the presence of added deoxynucleoside triphosphates, condense nucleotides to the 3' hydroxyl terminus in a single strand nick region of double-stranded DNA. Simultaneously the enzyme will remove nucleotides from

the 5' end of the nick. If one or more of the added triphosphates is labelled, for example, with α-^{32}P-phosphate, the label will be incorporated into the new strand by Pol I. Following the gap-filling procedure, recessed ends left after restriction endonuclease cutting can be filled in with Klenow fragments from Pol I or T4 DNA Polymerase.

Both the nick translation and the gap-filling procedures will yield double-stranded labelled DNA. The length of the product depends upon how much DNase I is added to the reaction. This is usually optimized so that thirty percent of the label is incorporated and the strands are 400 to 800 nucleotide units in length. The product length is unpredictably heterogeneous within the 400 to 800 unit range. In order to conserve enzyme as well as labelled nucleotide, only one microgram of DNA is usually labelled in each reaction vessel.

Double-stranded polynucleotides which incorporate pyrimidine bases modified by carbon chains at the C-5 position have been prepared enzymatically in a similar manner. This has been done by enzymatic elongation of a homopolymeric primer-template as reported by J. Sagi et al., Biochem. Biosphys. Acta. 606: 196-201 (1980). or by nick translation/gap filling by DNA polymerases using 2'-deoxyuridine 5'-triphosphate covalently linked to biotin as reported by P. Langer et al, Proc. Nat. Acad. Sci. USA 78:6633-6637 (1981), the biotin being capable of acting as a recognition site for avidin.

Enzymatic methods described by Rigby, et al, Bourguignon, et al, and Langer, et al, result in products having similar physical characteristics. Such enzymatically-prepared polynucleotides are 400-800 units in length,

require double-stranded polynucleotides as starting materials, produce double-stranded polynucleotides in all cases, and do not allow labeling at pre-selected sites.

In addition, all enzymatic methods modify both strands of the polynucleotide, and such product strands cannot be isolated from one another. During such processes, enzymes replace either all units with modified units or, when provided with a mixture of modified and naturally-occurring nucleoside triphosphates, randomly insert modified units. Furthermore, the enzymatic process described by Langer, et al, is incapable of producing polynucleotides incorporating such reporter groups as fluorescent, luminescent, or antigenic reporter groups. None of this art is capable of producing oligonucleotides of defined length, defined sequence or single-stranded character, either with or without reporter groups. Moreover, by the prior art methods, modified bases having reporter groups attached thereto cannot be incorporated in a polynucleotide at preselected sites.

Double-stranded polynucleotides which incorporate adenine bases modified at the C-8 position have also been prepared enzymatically. This has been done by incorporating 8-aminohexylamino-ATP (a ribonucleotide) onto DNA fragments, as reported by C. Vincent et al. Nucl. Acids Res. 10:6787-6796 (1982). The method is limited in scope, however, allowing only 3'-end labeling with triphosphates of adenine ribonucleotides. No modified pyrimidine nucleosides can be incorporated. Furthermore, as with other enzymatic methods, both strands of a double-stranded polynucleotide are labeled and no short (<100 units) oligonucleotides of defined sequence can be produced.

The prior art enzymatic methods referred to above require the chemical synthesis of a substituted nucleoside 5'-triphosphate, and demand subsequent enzymatic recognition and incorporation of the unnatural substrate into nicked double-stranded DNA. Such methods are incapable of producing polynucleotides of any preselected length or sequence, and the polymerases and DNases used therein are expensive. Moreover, these methods are time consuming, inefficient, demand substantial enzymatic activity and are limited to double-stranded DNA. Only small amounts, i.e., micrograms, of ill-defined polynucleotides, usually restriction fragments, are produced, and these must be tediously isolated from natural sources. Moreover, the scope of modifications obtainable in the polynucleotide product is severely limited, since the DNA polymerase cannot recognize or incorporate potentially useful reporter groups such as fluorescein or dinitrophenyl.

Attachment of one fluorescent molecule to the 3' end of long polyribonucleotide molecules (RNA) for limited biological application is disclosed by J.G.J. Bauman et al, J. Histochem. Cytochem. 29:238 (1981). This approach also used very small amounts (microgram quantities) of RNA tediously isolated from natural sources using enzymatic methodology, and cannot be applied to DNA since both 2' and 3' hydroxyls are required therefor. The much greater chemical instability of RNA relative to DNA also minimizes the scope of application of the polyribonucleotide produced by this method.

The non-enzymatic synthesis of defined sequence oligonucleotides incorporating naturally-occurring nucleic acid bases has been reported or reviewed by S.A. Narang et al, Meth. Enzymol 65:610 (1980), R. Letsinger, J. Org. Chem. 45:2715 (1980), M. Matteucci et al, J. Amer. Chem.

Soc. 103:3185 (1982), and G. Alvarado-Urbina et al, Science 214:270 (1981). Such synthesis usually involved chain extension by coupling an activated nucleotide monomer and a free hydroxyl-bearing terminal unit of a growing nucleotide chain. The coupling is effected through a phosphorus-containing group, as in the phosphate triester method reviewed by Narang et al, or one of the phosphite triester methods. Of the latter, those of Letsinger et al and Alvarado-Urbina et al use phosphochloridite chemistry, and that of Matteucci et al uses phosphoamidite chemistry.

The aforementioned chemical synthesis of oligonucleotides has incorporated only unmodified or naturally-occurring nucleic acid bases, and the end product there of is in short fragments and resembles unmodified or naturally-occurring RNA or DNA. It is also worthy of note that the end product of such synthesis does not incorporate any labels or reporter groups.

Direct modification of homopolymeric polynucleotides has been reported in systems of polyuridylic acid [Bigge, et al, J. Carb., Nucleosides, Nucleotides 8:259 (1981)]. The reported procedure is of limited scope and is not productive of useful products. Treatments described cause extensive polynucleotide cleavage and degradation, result in product irreversibly contaminated with metal ions, and are capable of modifying only cytosine residues of a DNA polynucleotide. The method is incapable of producing defined sequence oligonucleotides of specific lengths, cannot modify thymine or purine bases, and cannot modify at previously selected sites.

From the foregoing it will be apparent that labelled, defined sequence polynucleotides have been produced heretofor only by enzymatic methods which have the disadvantages pointed out herein, particularly those relating to time, cost, product length, sequence and yield. Moreover, such methodology is productive of only double-stranded products. Such double-stranded polynucleotides can be denatured by alkali or heat to cause spontaneous short-term separation of strands in solution. However, the individual single strands cannot be physically isolated from each other, and removal of the denaturing conditions results in rapid renaturation to double-stranded form. Since conditions productive of hybridization are also productive of renaturation, subjecting the denatured polynucleotide to hybridization conditions results in return of the polynucleotide to its original double-stranded configuration in which hybridization of either of the strands to a target polynucleotide is limited by competition from the other strand.

Modification of nucleosides has been undertaken, for example, in the synthesis of antiviral C-5 substituted pyrimidine nucleosides disclosed by Bergstrom, et al, J. Amer. Chem. Soc. 98:1587-1589 (1976); Ruth, et al, J. Org. Chem. 43:2870-2876 (1978); and Bergstrom, et al, U. S. Patent Nos. 4,247,544 and 4,267,171, or in the synthesis of C-8 substituted adenine derivatives disclosed by Zappelli, et al, U. S. Patent Nos. 4,336,188 and 4,199,498. These nucleosides, and others reported by Langer, et al, and D. Ward, European Patent Application No. 0063879, are not useful in the process of the present invention. Such reported nucleosides are highly reactive at undesired sites, and, if used in oligonucleotide synthesis by chemical methods, would result in undesired side products, uncontrollable synthesis, and no desired product. Moreover, such reported nucleosides do not contain proper sites in the substituent group, cannot be modified by the attachment of reporter groups, nor do they contain masked reactive functionalities. No such nucleosides are useful in the process of the present invention.

No chemical synthesis of defined sequence oligonucleotides incorporating modified bases of any kind, either with or without reporter groups, has been disclosed in the prior art.

There is an urgent need for high quality labelled, defined sequence single strand oligonucleotides which do not involve hazardous and unstable radioisotopes, and satisfaction of this need is a principal object of

the present invention. This object is accomplished by a chemical,. i.e., nonenzymatic, process which provides a predictable, superior product in high yield. More particularly, the process of the present invention accomplishes the chemical incorporation into defined sequence oligonucleotides of nucleotides modified with a wide variety of selected detectable reporter groups, such oligonucleotides being useful, for example, for the identification, localization, isolation and/or quantitation of complementary sequences of interest.

Another object of the invention is to provide novel nucleosides useful in the chemical synthesis of labelled, defined sequence single strand oligonucleotides.

The process of the present invention accomplishes the de novo chemical synthesis of labelled, defined sequence oligonucleotides and is superior to prior art enzymatic methods in a number of respects. More specifically the process of the present invention makes possible the synthesis of labelled, defined sequence single-stranded oligonucleotides of homogeneous predictably defined length, preferably having fewer than 200 base units, in contrast to the production of heterogeneous unpredictable populations of 400 to 10,000 base units of double-stranded character produced by prior art enzymatic methods.

The yield of product produced by the process of the present invention is of the order of hundreds to tens of thousands of micrograms, in contrast to the yield of a few micrograms provided by the prior art enzymatic methods. Moreover, the product oligonucleotides of the present invention are single-stranded, rather than the double-stranded products of enzymatic

methods. The single strand configuration of the product oligonucleotides avoids the competition from a complementary strand inherent in rehybridization of double-stranded polynucleotides.

Disclosure of Invention

The invention comprises a defined sequence single strand oligonucleotide of fewer than about 200 base units in length which comprises at least one nucleotide unit having attached to the base thereof at a sterically tolerant site (such as C-5 of pyrimidines and C-8 of purines) a substituent group which is capable of functioning as one or more detectable reporter groups, or of binding one or more detectable reporter groups. The invention also comprises a process for the chemical, i.e., nonenzymatic, synthesis of a defined sequence single strand oligonucleotide which comprises coupling an activated nucleotide monomer and a free hydroxyl-bearing terminal unit of a growing nucleotide chain, at least one of the monomer and terminal unit having its base modified at a sterically tolerant site by attachment thereto of a substituent group capable of functioning as one or more detectable reporter groups, or of binding one or more detectable reporter groups. The invention additionally comprises novel nucleosides useful in this synthetic process.

The oligonucleotide produced by the process of the present invention may include one or more pyrimidine-based or purine-based units which may be ribonucleotides or deoxyribonucleotides, and prior to the synthesis thereof, the reporter groups are preselected, as are the particular nucleotide units to which the reporter groups are attached.

Best Mode of Carrying Out the Invention

The chemical process by which the defined sequence single strand oligonucleotides of the present invention are preferably synthesized comprises coupling an activated nucleotide monomer and a free hydroxyl-bearing terminal unit of a growing nucleotide chain, at least one of said monomer and terminal unit having its base modified at a sterically tolerant site by attachment thereto of a substituent group capable of functioning as one or more detectable reporter groups, or of binding one or more detectable reporter groups.

The substituent groups of the present invention which are capable of binding reporter groups can be generally characterized as those which exhibit nucleophilic properties. Exemplary of such groups are those which contain primary or aromatic amines, carboxylic acids, hydroxyls and the like.

The bases of the nucleotide monomer and terminal unit are selected to provide the predetermined sequence of nucleotide units desired in the end product oligonucleotide. Such bases can take the form of the purines adenine (A), guanine (G), or hypoxanthine (H), or of the pyrimidines uracil (U), cytosine (C), or thymine (T). Such bases may also take the form of any other base which can be isolated from natural sources.

A sterically tolerant site on a nucleotide unit can be defined as a position on the nucleic acid base of the unit at which modification of the unit can be effected by attachment thereto of the substituent group without causing significant interference with hybridization of the product oligonucleotide and complementary nucleic acid components, and without sterically preventing the substituent group from functioning as one or more reporter groups, or from binding one or more reporter groups. Sterically tolerant sites are found at the C-8

position of purines and at the C-5 position of pyrimidines. Since the oligonucleotides of the present invention are particularly useful as hybridization probes, modifications which attach substituent and/or reporter groups should not be at sites on the pyrimidine or purine bases which are necessary for specific hybridization. Sites which should not be modified include N^1 and N^6 of adenine bases; N^1, N^2, and O^6 of guanine bases; N^3 and N^4 of cytosine bases. Generally, substitution at any heteroatom (N or O) should be avoided.

A reporter group can be defined as a chemical group which may be aromatic and/or polycyclic and which has a physical or chemical characteristic which can be readily measured or detected by appropriate physical or chemical detector systems or procedures. Reporter groups which are useful in oligonucleotides of the present invention are readily detectable. Ready detectability may be provided by such characteristics as color change, luminescence, fluorescence, or radioactivity; or it may be provided by the ability of the reporter group to serve as a ligand recognition site. Such groups are termed functionally colorimetric, luminescent, fluorescent, radioactive or ligand recognition groups. Among such groups are those suitable for ready detection by conventional detection techniques, such as colorimetric, spectrophotometric, fluorometric or radioactive detection, as well as those which are capable of participating inthe formation of specific ligand-ligand complexes which contain groups detectable by such conventional detection procedures.

A reporter group as used herein is defined as a substituent group which has physical, chemical or other

characteristics which can be readily measured or detected by the use of appropriate measurement or detection procedures. A reporter group as defined herein also includes a substituent group which is capable of initiating one or more interactions which ultimately provide a reaction product or a complex having physical, chemical or other characteristics which can be readily measured or detected by the use of appropriate measurement or detection procedures.

Exemplary of the measurable or detectable characteristics which such groups, reaction products or complexes may exhibit or induce are a color change, luminescence, fluorescence or radioactivity. Such characteristics can be measured or detected by the use of conventional colorimetric, spectrophotometric, fluorometric or radioactivity sensing instrumentation.

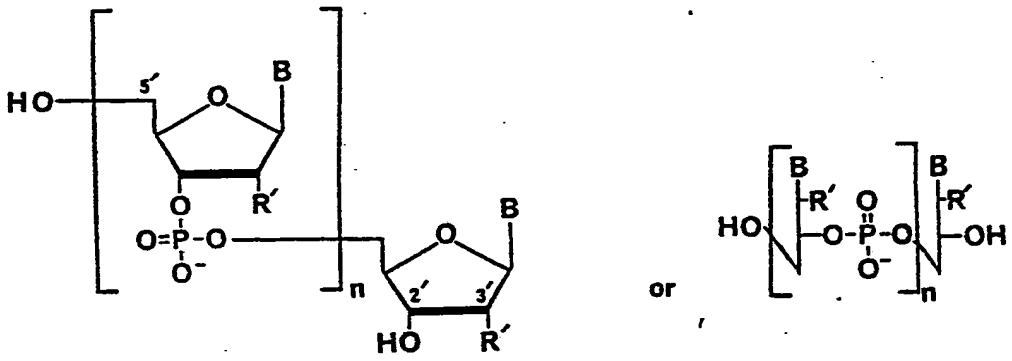
The interactions which usefully can be initiated by the reporter group defined herein include appropriately specific and selective ligand-ligand interactions productive of groups or complexes which are readily detectable, for example, by colorimetric, spectrophotometric, fluorometric, or radioactive detection procedures. Such interactions may take the form of protein-ligand, enzyme-substrate, antibody-antigen, carbohydrate-lectin, protein-cofactor, protein-effector, nucleic acid-nucleic acid or nucleic acid-ligand interactions. Exemplary of such ligand-ligand interactions are dinitrophenyl-dinitrophenyl antibody, biotin-avidin, oligonucleotide-complementary oligonucleotide, DNA-DNA, RNA-DNA and NADH-dehydrogenase. Other useful interactions will suggest themselves to those skilled in the art.

In the process of the present invention, a selected reporter group or groups can optionally be attached to the nucleotide monomer before coupling of the monomer to the terminal unit of the nucleotide chain, or it can be attached to the product oligonucleotide after formation thereof. The sequence of the nucleotide units in the product oligonucleotide is preselected to provide such oligonucleotide with precise specificity for its ultimate use. The oligonucleotides of the present invention are useful tools in recombinant DNA and other protocols involving nucleic acid rehybridization techniques. Among such uses are identification, localization, isolation and/or quantitation of complementary sequences of interest in cellular or cell-free systems. More specifically, such uses may include diagnostic applications or any fundamental biological event involving hybridization of nucleic acid components, or purification of complementary sequences by affinity chromatography when the product oligonucleotide is attached to a solid support through the modifications at a sterically tolerant site, with or without subsequent detection.

The nucleotide units in the product oligonucleotide may be purine or pyrimidine-based and may comprise units having naturally-occurring bases intermixed with units having modified bases. Such units may be ribonucleotides or deoxyribonucleotides. The coupling step preferably involves coupling of a monomer unit activated at the 3' position with a free 5' hydroxyl of the terminal unit of the growing nucleotide chain. Alternatively such coupling can involve coupling of a monomer unit activated at the 5' position with a free 3' hydroxyl of the terminal unit of the nucleotide chain. The

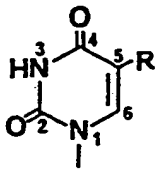
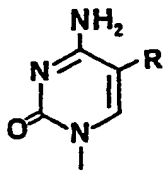
terminal unit may be the initial or only unit in the growing nucleotide chain at the time of coupling thereto of the nucleotide monomer, or it may be the terminal one of a plurality of nucleotide units.

The process of the present invention produces defined sequence oligonucleotides of the following generic formula:



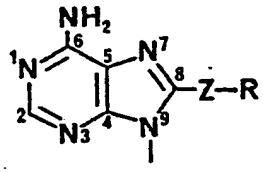
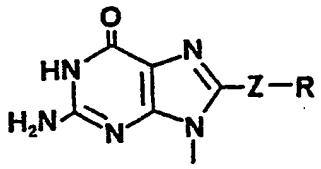
FORMULA I

Wherein n is 1 to about 199, preferably about 5 to about 60, and most preferably about 10 to about 40, R' is hydrogen or hydroxy, and B is any of the naturally-occurring purine or pyrimidine bases adenine, guanine, cytosine, uracil, thymine, or any other naturally-occurring base, the nucleotide units having naturally-occurring bases being independently intermixed with one or more nucleotide units having modified bases (B^m). The modified pyrimidine bases (Py^m) are substituted at the C-5 position, and typical examples thereof are the uracil and cytosine bases illustrated by the following generic formulas:

Py^m =  or 

modified uracil base modified cytosine base

The modified purine bases (Pu^m) are substituted at the C-8 position, and typical examples thereof are the modified adenine and guanine bases illustrated by the following generic formulas:

Pu^m =  or 

modified adenine base modified guanine base

The substituent group R is characterized by its ability to function as one or more reporter groups, or to bind one or more reporter groups. In the modified pyrimidine bases the substituent group R comprises two or more carbon atoms, whereas in the modified purine bases R comprises one or more carbon atoms. In this context,

R preferably takes the form of one of the following functionalized carbon chains:

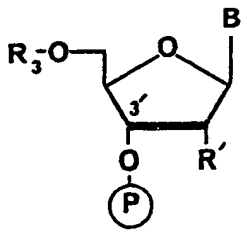
$$R = -CH{=}CR_1R_2,\ -CH_2CHR_1R_2,\ -CHR_1R_2,$$

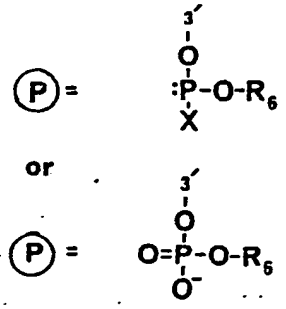
$$-CH{=}CR_1-\overset{O}{\overset{\|}{C}}NHR_2,\ \text{or}\ -CH{=}CR_1-\overset{O}{\overset{\|}{C}}-R_2$$

wherein R_1 is hydrogen or alkyl; R_2 is alkyl, alkenyl, aryl, or functionalized alkyl, alkenyl, aryl wherein functional groups include one or more amines, amides, nitriles, carboxylic acids and esters, hydroxy, dinitrophenyl, aminobenzenesulfonates, or the like; and Z is a polyvalent heteroatom such as nitrogen, oxygen or sulfur. In addition, R_2 may be attached to a solid support, or to one or more reporter groups which function, for example, as a colorimetric, fluorescent, luminescent, radioactive, or ligand recognition group. Functionally fluorescent groups include fluoresceins, rhodamines, and the like or adducts thereof; functionally luminescent groups include luminols, acridines, luciferins, dioxetanes, dioxamides, and the like or adducts thereof. Ligand recognition groups include vitamins (such as biotin, iminobiotin or desthiobiotin), antigens such as dinitrophenols, carbohydrates and other functional groups or adducts of such groups which can be recognized by ligand-like interactions with proteins, or from which such ligand-like interactions can be elicited. Another oligonucleotide capable of interaction with nucleic acids is illustrative of a group from which a ligand-like interaction can be elicited. Ligand recognition groups may also serve as functionally colorimetric reporter groups when recognition results in dye formation. For example, when dinitrophenyl is used as a reporter group, known detection systems using an antidinitrophenyl antibody coupled to peroxidase can be used as a detection system, resulting in a color change. Functionally radioactive groups incorporate a radioactive element in the chosen reporter group.

When reference is made herein to the use of purine or pyrimidine bases, such expressions are intended to include analogs of such bases. Among such analogs are the analogs of purine bases, such as the deazaadenosines (tubercidins, formycins, and the like), and the analogs of pyrimidine bases, such as deazauracil, deazacytosine, azauracils, azacytosines, and the like.

Oligonucleotides of Formula I are best prepared by chemical synthesis from monomer nucleotide analog units of the formula:





where:

R_6 = methyl, chlorophenyl

X = chloro, dialkylamino, morpholino

FORMULA II

wherein R_3 is trityl (triphenylmethyl), dimethoxytrityl, or other appropriate masking group for the 5'-hydroxyl; B and R' are masked, if appropriate; and (P) represents a phosphorus-containing group suitable for internucleotide bond formation during chain extension in synthesis of a product oligonucleotide. The phosphorus-containing groups (P) suitable for internucleotide bond formation are preferably alkyl phosphomonochloridites or alkyl phosphomonoamidites. Alternatively phosphate triesters may be employed for this purpose. The monomer unit may

alternatively have R_3 attached at the 3' hydroxyl and Ⓟ attached at the 5'-hydroxyl.

Generally, the term "masking group" or "blocking group" is a functional expression referring to the chemical modification or "blocking" of an integral functional group by attachment of a second moiety to disguise the chemical reactivity of the functional group and prevent it from reacting in an undesired manner. Such modification is reversible, and allows subsequent conversion back to the original functional group by suitable treatment. In many cases, such masking formally interconverts structural functionality, e.g., a primary amine masked by acetylation becomes a substituted amide which can be later converted back to the primary amine by appropriate hydrolysis.

The compounds of Formula I include the acceptable conjugate acid salts thereof. Conjugate acids which may be used to prepare such salts are those containing nonreactive cations and include, for example, nitrogen-containing bases such as ammonium salts, mono-, di-, tri- or tetra-substituted amine salts, and the like, or suitable metal salts such as those of sodium, potassium, and the like.

The process steps of the present invention will now be generally described and illustrated diagrammatically. Thereafter, the invention will be illustrated more specifically and detailed examples therof provided. Since the invention relates to oligonucleotides incorporating both pyrimidine-based and purine-based nucleotide units, the use of both pyrimidine and purine-based compounds in the synthetic process will be illustrated. The specific pyrimidine and purine-based compounds illustrated are only exemplary of the respective pyrimidine and purine classes, and it is to be understood that any other member

of the respective class can be substituted therefore in the process and the product oligonucleotide, whenever suitable or desired. While deoxyribonucleotide compounds are shown for the most part, it is understood that ribonucleotide compounds are also contemplated by the invention and can be substituted for the deoxyribonucleotide compounds wherever ribonucleotide compounds are desired in the product oligonucleotide.

One of the more important aspects of the invention is the provision of a new class of nucleosides which are essential as intermediates in the process for synthesizing the new oligonucleotides. Such nucleosides each have a base which is modified by a substituent group comprising a functionalized carbon chain and one or more amides, the nitrogen of the amides being attached to a sterically tolerant site on the base through the carbon chain. In the case of pyrimidine-based nucleosides, the carbon chain is attached at the C-5 position, and in the case of the purine-based nucleosides, the carbon chain is attached at the C-8 position through a polyvalent heteroatom, such as nitrogen, oxygen or sulfur. In addition, such nucleosides are chemically blocked at the 5' position (or the 3' position) with a group, such as dimethoxytrityl, appropriate for the chemical synthesis of oligonucleotides.

In the new class of nucleosides the substituent group may be chosen from $-CH_2CHR_1C_nH_{2n}Y$, $-CH=CR_1C_nH_{2n}Y$, $-CH=CR_1-\overset{O}{\overset{\|}{C}}-NHC_nH_{2n}Y$, or $-CH=CR_1\overset{O}{\overset{\|}{C}}C_nH_{2n}Y$ wherein R_1 is hydrogen or C_{1-6} lower alkyl and Y is one or more amido, substituted amido, substituted amino or substituted aminoalkylphenyl groups. More specifically, Y may include one or more $-NH\overset{O}{\overset{\|}{C}}CX_3$ wherein X is hydrogen, fluorine or chlorine. Synthesis of these nucleosides, as well as of the masked forms thereof, is described hereinafter in Examples I, II, IV, VI, VII, VIII, X, XII, and XIII. Preferred nucleosides incorporate the substituent group $-CH=CH\overset{O}{\overset{\|}{C}}-NHC_nH_{2n}Y$ at C-5 of pyrimidine nucleosides wherein n = 3 to 12 and Y is $-NH\overset{O}{\overset{\|}{C}}CX_3$. Most preferred are such nucleosides wherein the pyrimidine base is uracil.

The process of the present invention may be initiated by the preparation of the selected nucleoside. Generally, the most preferred nucleosides are best prepared in the following manner. 5-(Methyl 3-acrylyl) -2'-deoxyuridine is prepared from 2'-deoxyuridine by the method of Bergstrom and Ruth [J. Amer. Chem. Soc. 96:1587 (1976)]. The nucleoside is then treated with 1.05 equivalents of dimethoxytrityl chloride in pyridine for 4 hours to block the 5'-hydroxyl with dimethoxytrityl (DMT). The resulting product is purified by silica chromatography eluting a gradient of 0-10% methanol in chloroform containing 2% triethylamine. The purified 5'-DMT-5-(methyl 3-acrylyl)-2'-deoxyuridine is treated with 1 N KOH for 24 hr. at ambient temperature to hydrolyze the methyl ester.

The resulting 5'-DMT-5-(3-acrylyl)-2'-deoxyuridine is treated with excess dicyclohexylcarbodiimide and hydroxybenztriazole in pyridine. After 4 hours, a 2-5 fold excess of 1,7-diaminoheptane is added, and the reaction stirred overnight. After 12-20 hours, a 10-20 fold excess of trifluoroacetic anhydride is added, and the reaction stirred at room temperature for 4 hours. The product is purified by silica chromatography eluting a gradient of 0-10% methanol in chloroform containing 2% triethylamine, followed by exclusion chromatography using Sephadex LH-20 eluting 1% triethylamine in methanol. Appropriate fractions are combined to yield 5'DMT-5-[N-(7-trifluoroacetylaminoheptyl)-1-acrylamido]-2'-deoxyuridine; such product is appropriate for oligonucleotide synthesis by the phosphochloridite procedure described in Examples XV and XVIII. Alternatively, such a compound can be prepared by the combination of methods described in Examples II and III. Replacing diaminoheptane in this process with other diaminoalkanes (e.g., diaminopropane, diaminohexane, diaminododecane) is productive of other compounds of varying substituent length wherein n = 3, 6, or 12 and R = $-CH=CHC(=O)-NHC_nH_{2n}NHC(=O)CX_3$. Two such nucleosides, one pyrimidine (uracil)-based and the other purine (adenine)-based, are shown at the top of the diagram below illustrating the process. Reactive sites on the bases of the nucleosides are then masked, as shown in Reaction 1, by attachment of, for example, a benzoyl group (Bz) to the amine at the

6 position of the adenine-based nucleoside. Such masking is generally described in "Synthetic Procedures in Nucleic Acid Chemistry", Vol. 1, W. Zorbach and R. Tipson eds. (Wiley - Interscience, N. Y.) 1968). Unprotected amines on the substituent group are masked, for example, by attachment thereto of trifluoroacetyl groups (Ac), as also shown in Reaction 1.

The selected 3' or 5' hydroxyl of the nucleoside is then masked by attachment thereto of a dimethoxytrityl (DMT) group. In Reaction 2 illustrated below, the 5'-hydroxyl is masked, leaving the 3' hydroxyl free or available for reaction. Alternatively, the 3' hydroxyl could be masked, leaving the 5' hydroxyl free.

The nucleoside is then converted to an activated nucleotide monomer, preferably by attachement to its 3' hydroxyl of a phosphorus-containing group which includes an activating moiety. When the modified nucleoside is properly blocked, modifications of the procedures described by Letsinger, et al, Matteucci, et al, or as reviewed by Narang, et al can be utilized for oligonucleotide synthesis. The use of phosphochloridite chemistry such as that disclosed by Letsinger, et al, is detailed in Examples XVI-XVIII. In order to use phosphoamidite chemistry, a modification of the procedure of Matteucci, et al, is used, phosphitylating the protected modified nucleoside with methyl chloro (N, N-diisopropyl)phosphoamidate or methyl chlorophosphomorpholidate, as in the improved procedure of Dorper, et al [Nucleic Acids Res. 11:2575(1983)]. Al-

ternatively, the protected modified nucleoside can be phosphorylated with 1.2 eq. chlorophenyl dichlorophosphate in trimethylphosphate at room temperature followed by quenching with water to give the 3'-chlorophenyl phosphate adduct of the modified nucleoside, such adducts being useful in a modification of the phosphotriester approach as illustratively reviewed by Narang, et al. The diagram illustrates in Reaction 3 the synthesis of activated monomer nucleotide units of Formula II by attachment to the nucleoside 3' hydroxyl of a phosphomonochloridite group in which the chlorine functions as an activating moiety.

Coupling or condensation of the selected activated nucleotide monomer, i.e. the uracil-based monomer or the adenine-based monomer, to the terminal unit of a growing nucleotide chain is illustrated in Reaction 4 in the diagram. The nucleotide chain is shown as including in its right hand end a nucleoside unit having a naturally occurring base and having a solid support or masking group R_4 attached to its 3' hydroxyl. The illustrated chain also includes one or more (n') nucleotide units having naturally-occurring bases, said units being coupled to the 5' hydroxyl of the nucleoside unit, the terminal one of the nucleotide units having a free hydroxyl at the 5' position. In the coupling reaction

This Page is inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☒ BLURED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLORED OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REPERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.
As rescanning documents *will not* correct images problems checked, please do not report the problems to the IFW Image Problem Mailbox